RAW SEQUENCE LISTING
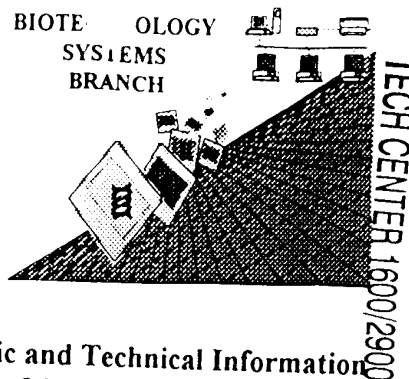# ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _O 9 5/8/5 6_

Source: _AU 1645_

Date Processed by STIC: _08/06/2001_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE __CHECKER VERSION 3.0 PROGRAM__, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1 821 - 1 825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

Does Not Comply   **1645**
Corrected Diskette Needed

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/518,156   DATE: 08/06/2001
                                              TIME: 15:04:53

Input Set : **A:\sequence.listing.txt**
Output Set: **N:\CRF3\08062001\I518156.raw**

```
 3 <110> APPLICANT: Tarleton, Rick
 4        Garg, Nisha
 6 <120> TITLE OF INVENTION: PROPHYLACTIC AND THERAPEUTIC IMMUNIZATION AGAINST INFECTION
AND DISEASE
 8 <130> FILE REFERENCE: 235.00201010
10 <140> CURRENT APPLICATION NUMBER: 09/518,156
11 <141> CURRENT FILING DATE: 2000-03-02
13 <150> PRIOR APPLICATION NUMBER: 60/122,532
14 <151> PRIOR FILING DATE: 1999-03-02
16 <160> NUMBER OF SEQ ID NOS: 24
18 <170> SOFTWARE: PatentIn version 3.0
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 8
22 <212> TYPE: PRT
23 <213> ORGANISM: Trypanosoma cruzi
25 <400> SEQUENCE: 1
27 Val Asp Tyr Asn Phe Thr Ile Val
28 1               5
30 <210> SEQ ID NO: 2
31 <211> LENGTH: 8
32 <212> TYPE: PRT
33 <213> ORGANISM: Gallus gallus
35 <400> SEQUENCE: 2
37 Ser Ile Ile Asn Phe Glu Lys Leu
38 1               5
40 <210> SEQ ID NO: 3
41 <211> LENGTH: 34
42 <212> TYPE: DNA
43 <213> ORGANISM: Artificial
45 <220> FEATURE:
46 <223> OTHER INFORMATION: forward primer
48 <400> SEQUENCE: 3
49 agtcgacgga tccatgattg catttgtcga aggc                           34
52 <210> SEQ ID NO: 4
53 <211> LENGTH: 35
54 <212> TYPE: DNA
55 <213> ORGANISM: Artificial
57 <220> FEATURE:
58 <223> OTHER INFORMATION: reverse primer
60 <400> SEQUENCE: 4
61 atctagaagc ttcatagttc accgacactc agtgg                          35
64 <210> SEQ ID NO: 5
65 <211> LENGTH: 35
66 <212> TYPE: DNA
67 <213> ORGANISM: Artificial
69 <220> FEATURE:
70 <223> OTHER INFORMATION: reverse primer
72 <400> SEQUENCE: 5
```

C--> 43 **<213> ORGANISM: Artificial**

C--> 55 **<213> ORGANISM: Artificial**

C--> 67 **<213> ORGANISM: Artificial**

```
    73 atctagaagc ttcatgccgc agcatttgct tcccc                                    35
    76 <210> SEQ ID NO: 6.
    77 <211> LENGTH: 0
    78 <212> TYPE: PRT
    79 <213> ORGANISM: Skipped Sequence )
    81 <400> SEQUENCE: 6
W--> 82 000
    84 <210> SEQ ID NO: 7
    85 <211> LENGTH: 8
    86 <212> TYPE: PRT
    87 <213> ORGANISM: Trypanosoma cruzi
    89 <400> SEQUENCE: 7
    91 Val Asn His Arg Phe Thr Leu Val
    92 1               5
    94 <210> SEQ ID NO: 8
    95 <211> LENGTH: 8
    96 <212> TYPE: PRT
    97 <213> ORGANISM: Trypanosoma cruzi
    99 <400> SEQUENCE: 8
   101 Val Asn His Asp Phe Thr Val Val
   102 1               5
   104 <210> SEQ ID NO: 9
   105 <211> LENGTH: 0
   106 <212> TYPE: PRT
   107 <213> ORGANISM: Skipped Sequence
   109 <400> SEQUENCE: 9
W--> 110 000
   112 <210> SEQ ID NO: 10
   113 <211> LENGTH: 28
   114 <212> TYPE: DNA
C--> 115 <213> ORGANISM: Artificial
   117 <220> FEATURE:
   118 <223> OTHER INFORMATION: forward primer
   120 <400> SEQUENCE: 10
   121 aggatccatg attgcatttg tcgaaggc                                            28
   124 <210> SEQ ID NO: 11
   125 <211> LENGTH: 30
   126 <212> TYPE: DNA
C--> 127 <213> ORGANISM: Artificial
   129 <220> FEATURE:
   130 <223> OTHER INFORMATION: reverse primer
   132 <400> SEQUENCE: 11
   133 aaagcttcat agttcaccga cactcagtgg                                          30
   136 <210> SEQ ID NO: 12
   137 <211> LENGTH: 26
   138 <212> TYPE: DNA
C--> 139 <213> ORGANISM: Artificial
   141 <220> FEATURE:
   142 <223> OTHER INFORMATION: forward primer
```

errored

errored
_When a sequence is skipped
only fields 210 and 400 are
included. (NW)_

The types of errors shown exist _throughout_ the Sequence Listing. Please check subsequent sequences for similar errors.

```
      144 <400> SEQUENCE: 12
      145 aagatcttgt ggaaaggaat ttgagg
      148 <210> SEQ ID NO: 13                                         26
      149 <211> LENGTH: 28
      150 <212> TYPE: DNA
C--> 151 <213> ORGANISM: Artificial
      153 <220> FEATURE:
      154 <223> OTHER INFORMATION: reverse primer
      156 <400> SEQUENCE: 13
      157 actcgagtca cagtgggcgg ttgtacag
      160 <210> SEQ ID NO: 14                                         28
      161 <211> LENGTH: 27
      162 <212> TYPE: DNA
C--> 163 <213> ORGANISM: Artificial
      165 <220> FEATURE:
      166 <223> OTHER INFORMATION: forward primer
      168 <400> SEQUENCE: 14
      169 aagatctctg tgaggctgca gacgctg
      172 <210> SEQ ID NO: 15                                         27
      173 <211> LENGTH: 28
      174 <212> TYPE: DNA
C--> 175 <213> ORGANISM: Artificial
      177 <220> FEATURE:
      178 <223> OTHER INFORMATION: reverse primer
      180 <400> SEQUENCE: 15
      181 acccgggtta ttggtcgcca ccgtttcc
      184 <210> SEQ ID NO: 16                                         28
      185 <211> LENGTH: 26
      186 <212> TYPE: DNA
C--> 187 <213> ORGANISM: Artificial
      189 <220> FEATURE:
      190 <223> OTHER INFORMATION: forward primer
      192 <400> SEQUENCE: 16
      193 ggttcgattg gggttggtgt aatata
      196 <210> SEQ ID NO: 17                                         26
      197 <211> LENGTH: 26
      198 <212> TYPE: DNA
C--> 199 <213> ORGANISM: Artificial
      201 <220> FEATURE:
      202 <223> OTHER INFORMATION: reverse primer
      204 <400> SEQUENCE: 17
      205 aaataatgta cgggkgagat gcatga
      208 <210> SEQ ID NO: 18                                         26
      209 <211> LENGTH: 0
      210 <212> TYPE: DNA
      211 <213> ORGANISM: Skipped Sequence
      213 <400> SEQUENCE: 18
W--> 214 000
      216 <210> SEQ ID NO: 19
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/518,156

DATE: 08/06/2001
TIME: 15:04:53

Input Set : A:\sequence.listing.txt
Output Set: N:\CRF3\08062001\I518156.raw

```
217 <211> LENGTH: 0
218 <212> TYPE: DNA
219 <213> ORGANISM: Skipped Sequence
221 <400> SEQUENCE: 19
222 000
224 <210> SEQ ID NO: 20
225 <211> LENGTH: 1656
226 <212> TYPE: DNA
227 <213> ORGANISM: Trypanosoma cruzi
229 <400> SEQUENCE: 20
230 atgcggaaga aagccgcagc attagcagcg cccacagcag acacacggcc gacgtgccgc      60
232 ggggctgcca ttgccaataa atttatggaa cgtgccggcc cccgtgaggg cgttgggaga     120
234 tcaacggaga tgccggctgc tggaccgacg gggtctcaaa gaactcaaac gcaacgggag     180
236 gtgaaagcgt cacaagacgc cgacgcggcg gccattagta gttatttcca gtccgaattg     240
238 gtgacatctc agtcgcacga gggtgtgtct cctctggcaa agactagggc caacgaacgg     300
240 cggaacgggg agcaggagcg ggagaaggaa ctgccggcgg ttggtggcgc cgttccaact     360
242 gggaagggga cggaccccaa acagcgagtg ctgcaggatt tgccagcgat gcacgcggag     420
244 ggacaaaacc agcacggtag agagggtgac aagggtgttt ccgtgaagat ggactcccct     480
246 ggtcgcgtac aggtgctgga gcaaatgttg ctacacctgg ctgcattgaa cagacagcta     540
248 gaattagaac ttatagaaac gcgacgggaa ctgacgatgt acaagcagct tttacctgat     600
250 gtgcagcgcc agaccgaggc ccatgctttg tctcaggagc atcacaaagc gaatagtgct     660
252 gctccgccac tgatgtcaga tgagaggcga cgacagatgc tctttacagg gcaacaacaa     720
254 caacagcaac aagtggaaga tctgcatggc ggtattagcg ggtgggaaac ggcagcgagg     780
256 agaatgcgct atggttacga ggaggggggag agggacgccc tttcagatgg tgagggccgt     840
258 ccacgttgcg caggtcgtat gggctccccg aagagattcc tttcaacaca accgcctcga     900
260 agcagcagga accatcggaa ccctcacgct gctaacggga caaatggcaa tagtcatgtt     960
262 ccccattcgt ccagacaaaa aagtcacccg acaagaggag ctgctgtaac ttccgtaccg    1020
264 ttggcggcgt ccgcaaccaa tcgccgaggt cgttccatgc gacaacatac ccgaccccgc    1080
266 ggaccttctt atcttttcga acgcctcgac gctgaggatg caattgatat gctggagacg    1140
268 ctgaagcgct ctctcatgta tcgctgcaac cactcgcatc atcgatcaac agaaggagat    1200
270 gttgtgcggc ccgccgcgaa gccccggaaa ggcacgcggt ctgttccacc accaccgcca    1260
272 ccaccgccca tgtcatcatc gtcacaaaga aagcttgccg ccgcagttgc tggagcgccg    1320
274 gcatgcagcg tctcagcacg acacggaagg aaccatggcg tttctgcggt gggagatccg    1380
276 tcaaggggca atcgagtttc agaaacagct cgcatagctc atgctccttc ttttgggggg    1440
278 aagaaatgcg cgccgggcct aacccaactc catttctctt ccccttccag aagggctacg    1500
280 ccgatgaaaa aagacacgcc attgtcacgt ggtcaagcgg ctggagtagc agcagtagcg    1560
282 gtgggcggtg acgggcagct agaggcactg cagaggcgtt actgggaaca gtcccgtgcg    1620
284 atattggagc agcttgaaaa catgctggca gctgat                             1656
287 <210> SEQ ID NO: 21
288 <211> LENGTH: 39
289 <212> TYPE: DNA
290 <213> ORGANISM: Artificial
292 <220> FEATURE:
293 <223> OTHER INFORMATION: pTAT linker
295 <400> SEQUENCE: 21
296 tccaccatgg ccggtaccgg tctcgaggtg catgcggtg                           39
299 <210> SEQ ID NO: 22
300 <211> LENGTH: 14
301 <212> TYPE: PRT
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/518,156   DATE: 08/06/2001
                                            TIME: 15:04:53
        Input Set : A:\sequence.listing.txt
        Output Set: N:\CRF3\08062001\I518156.raw

```
C--> 302 <213> ORGANISM: Artificial
     304 <220> FEATURE:
     305 <223> OTHER INFORMATION: pTAT linker
     307 <400> SEQUENCE: 22
     309 Gly Ser Thr Met Ala Gly Thr Gly Leu Glu Val His Ala Val
     310 1               5                   10
     312 <210> SEQ ID NO: 23
     313 <211> LENGTH: 41
     314 <212> TYPE: DNA
C--> 315 <213> ORGANISM: Artificial
     317 <220> FEATURE:
     318 <223> OTHER INFORMATION: pTAT-HA linker
     320 <400> SEQUENCE: 23
     321 ccatgtccgg ctatccatat gacgtcccag actatgctgg c            41
     324 <210> SEQ ID NO: 24
     325 <211> LENGTH: 13
     326 <212> TYPE: PRT
C--> 327 <213> ORGANISM: Artificial
     329 <220> FEATURE:
     330 <223> OTHER INFORMATION: pTAT-HA linker     good
     332 <400> SEQUENCE: 24
     334 Met Ser Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Gly
     335 1               5                   10
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/518,156   DATE: 08/06/2001
TIME: 15:04:54

Input Set : **A:\sequence.listing.txt**
Output Set: **N:\CRF3\08062001\I518156.raw**

```
L:43  M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:3
L:55  M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:4
L:67  M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5
L:82  M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (6) SEQUENCE:
L:110 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (9) SEQUENCE:
L:115 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:10
L:127 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:11
L:139 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:12
L:151 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:13
L:163 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:14
L:175 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15
L:187 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:16
L:199 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:17
L:214 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (18) SEQUENCE:
L:222 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (19) SEQUENCE:
L:290 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:21
L:302 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:22
L:315 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:23
L:327 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:24
```

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 09/513,196

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1_____Wrapped Nucleics / Wrapped Aminos — The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2_____Invalid Line Length — The rules require that a line not exceed 72 characters in length. This includes white spaces.

3_____Misaligned Amino Numbering — The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4_____Non-ASCII — The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5_____Variable Length — Sequence(s)____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6_____PatentIn 2.0 "bug" — A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7_____Skipped Sequences (OLD RULES) — Sequence(s)_____missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8_____Skipped Sequences (NEW RULES) — Sequence(s)_____missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9_____Use of n's or Xaa's (NEW RULES) — Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10_____Invalid <213> Response — Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11_____Use of <220> — Sequence(s)_____missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12_____PatentIn 2.0 "bug" — Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13_____Misuse of n — n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001